Title
On the Genesis of Organizational Forms: Evidence from the Market for Disk Arrays

Authors
David G. McKendrick
Glenn R. Carroll

Publication Date
2000-04-01

ON THE GENESIS OF ORGANIZATIONAL FORMS: EVIDENCE FROM THE MARKET FOR DISK ARRAYS

David G. McKendrick
Information Storage Industry Center
University of California at San Diego
La Jolla, CA 92093-0519
dmckendrick@ucsd.edu

Glenn R. Carroll
Haas School of Business
University of California at Berkeley
Berkeley, CA 94720
carroll@haas.berkeley.edu

Report 2000-01 January, 2000

(Revised April 2000)

The Information Storage Industry Center

Graduate School of International Relations and Pacific Studies
University of California
9500 Gilman Drive
La Jolla, CA 92093-0519
http://www-irps.ucsd.edu/~sloan/

University of California, San Diego

Funding for the Information Storage Industry Center
is provided by the Alfred P. Sloan Foundation

On the Genesis of Organizational Forms: Evidence from the Market for Disk Arrays

David G. McKendrick
Information Storage Industry Center
University of California at San Diego
La Jolla, CA 92093-0519
858-534-9824
dmckendrick@ucsd.edu

Glenn R. Carroll
Haas School of Business
University of California at Berkeley
Berkeley, CA 94720
510-642-0829
carroll@haas.berkeley.edu

April 2000

On the Genesis of Organizational Forms: Evidence from the Market for Disk Drive Arrays*

David G. McKendrick and Glenn R. Carroll
dmckendrick@ucsd.edu and carroll@haas.berkeley.edu

Abstract

This paper asks a basic question of organizational evolution: When and where will a new organizational form emerge? Contemporary organization theory proposes two answers. The first holds that formal institutions such as industry associations and standard-setting bodies will result in a taken-for-granted organizational form. The second answer contends that increasing organizational density (number of organizations in a population) will generate a legitimated organizational form.

Our detailed historical case study of the disk array market and its associated technologies suggests each of these theoretical arguments is limited. Although we find significant collective activity in association-building and standard-setting among disk array producers, these have not yet led to an organizational form. Similarly, an observed trajectory of organizational density showing rapid growth followed by stabilization has not yet generated an organizational form. In our view, the diversity of origins and other activities of those organizations operating in this market contribute to the lack of institutionalization of the disk array organizational form. We reason that if firms in the market derive their primary identities from other activities (implying that there are few highly focused firms deriving their primary identity from disk arrays), then the disk array producer identity cannot cohere into a code or form. This observation suggests a respecification of the legitimation component of the density-dependent model of organizational evolution.

*This research was supported by assistance from the Alfred P. Sloan Foundation and the Institute of Industrial Relations, U.C. Berkeley. We appreciate the helpful comments of Michael T. Hannan on an earlier draft and the research assistance of Yuko Kasuya.

Organizational theorists and researchers use the organizational form concept in many ways, often without precise definition. This paper builds on a recent attempt to define and theorize about organizational forms in rigorous fashion, using the tools of logical formalization (Polos et al., 1998a; 1998b; 1999; Carroll and Hannan 2000). It does so by tracing the possible emergence and development of a potentially new organizational form. The goal is to gain insight into the processes that generate and establish organizational forms and, ultimately, to develop systematic theory about those processes. We also hope to sharpen understanding of differences among various uses of the form concept.

Our research strategy is unconventional. It involves identifying and documenting a product--disk drive arrays--for which there is a large market but which may or may not spawn an institutionalized organizational form by our definition. In other words, we chose a *potential* organizational form in midstream. We did so to catch the form-generating processes in action and to minimize retrospective bias problems that would arise by choosing an established form. Whether the product and its market will eventually produce an organizational form or not is beside the point: we do not regard this effort as an exercise in theory testing. Rather, we believe that examining in detail the "facts" surrounding an emergent organizational form should be highly informative and useful for generating ideas and developing theory, regardless of the eventual outcome.

We expect that many readers will be unfamiliar with disk drive arrays (hereafter, disk arrays). That is not surprising since one of the things associated with institutionalized organizational forms is widespread awareness of what organizations with the form produce and do (Hannan and Carroll 1992; Scott 1995). As we explain in detail below, disk arrays are a particular type of technological solution to data storage problems. Broadly defined, disk arrays consist of multiple hard disk drives linked by proprietary software or hardware (in the form of a controller) that protects the integrity of stored data. Customers find disk arrays attractive because they provide reliable online access to large databases at low cost.

Disk arrays are widely used: industry data show that over $12.6 billion of disk array products were sold in 1998 (see Table 1). By our count, 130 producers of disk array products were operating at the end of 1998. However, disk arrays are but one of several major technological options facing those in the data storage market, including optical drives, hard drives, and tapes. Implicit in our research is the notion that sustained long-term viability of disk array producers depends not only on the technological features of array products but also on establishment of the array organizational form.

(Table 1 about here)

The paper is structured as follows. In the next section, we review and explain the Polos et al. (1998a; 1998b; 1999) definition of organizational form. We also discuss the types of processes that organizational theorists suggest might be behind the emergence and development of an organizational form. The following section briefly describes the technology and explains why we do not think that its producers hold a legitimated organizational form. We then review the early history of disk arrays with an eye to the origins of the concept and product. Next we describe the various industry associations that have been established in this area. Following that, we move to a description of the firms that have been involved in disk array production, looking at entry and exit patterns. Then, we discuss again the relevant theories in light of the descriptive evidence. This exercise suggests the theories have some limitations and leads us to advance a new argument, one that differs from other current efforts. We conclude by discussing the implications of our research.

ORGANIZATIONAL FORMS: CONCEPTS, DEFINITIONS AND THEORY

As explained by Polos et al. (1998a), the organizational form concept plays three major roles in sociological theory and research (see also Carroll and Hannan 2000). First, researchers use notions of form to define populations of organizations for study. Second, form refers to a selection-favored conglomerate of properties, often embodied in a structural architecture. Third, the form concept seeks to differentiate between core and peripheral features.

These uses of the form concept rely on several different underlying definitions of form. Perhaps the most common definition is based on features of organizations; others rely on boundary generating notions, including network boundaries. The most serious limitation of all current definitions, Polos et al. (1998a) note, is their failure to link forms and identities (see also Ruef 1999a; 1999b). Organizational forms presumably involve social and cultural typifications, widely agreed-upon classifications of entities into types. Efforts to understand such typifications in empirical research on diverse kinds of organizations suggest that such processes build upon identities. So, Polos et al. (1999) and Ruef (1999b) sought to define identity and form in closely related terms.

Polos et al. (1999) proposed a new theoretical approach to organizational forms, one that bases a system of classification of forms on a theory of action.[1] They view a form as a recognizable pattern that takes on rule-like standing, which they call a code. The term code denotes and connotes both cognitive recognition and imperative standing. A code can be understood as (1) a set of signals, as in the "genetic code" and (2) as a set of rules of conduct, as in the "penal code." Use of the term code by Polos et al. (1999) reflects both meanings. By their definition, a form is an external identity code that possesses rule-status, meaning that its observable violation causes outsiders to drop discontinuously their valuation of the entity to which it is applied.[2] Identity codes may consist of abstract features as well as composition rules that tell about appropriate combinations of features. Forms apply to multiple entities and persist better than ordinary identities.

Forms can be graded in terms of their degrees of specificity or sharpness. Forms might also be hierarchically nested with respect to each other. If so, then systems of forms may appear as semi-lattices or trees or diamonds or many other structures. These are questions of structural shape in the systems of forms. The degree of specificity in form-to-form relations is also an important issue, whatever the shape of a system. For example, a tree-shaped system might contain few vertical levels but have wide horizontal branches (wide spans). Conversely, a tree form might be narrow at each level but deep vertically.

Why does the institutional classification of an organizational form matter? As the anthropologist Douglas (1986: 103) notes about a similar classification system for persons: "Something happens to the insides of our heads when a different kind of organization had made obsolete the old classifications…The change is not a deliberate or conscious choice. Institutions veil their influence, so that we hardly notice any change." For classification of organizational forms, virtually *all* contemporary theories assign a privileged status to institutionalized (as opposed to uninstitutionalized) organizational forms. The returns to institutionalization include ease of resource mobilization and survival advantages. For example, Hannan and Carroll (1992: 36-7) argue:

> …a taken for granted social form can be more readily visualized by potential organizers than one with dubious or unknown standing. Variations in the strength of institutional rules endorsing rational organization as the appropriate vehicle for attaining collective goals affect the ease of founding organizations. The capacity to mobilize potential members and resources increases greatly when those who control resources take the organizational form for granted. Reducing the need for such justifications lowers the cost of organizing…Legitimation eases the problem of maintaining flows of resources from the environment and enhances the ability of organizations to fend off challenges.

A business illustration of the operation of these forces can be found in a recent account in the Wall Street Journal about Reuters, a media company founded in the 19th century and now engaged in a variety of internet businesses as well (Goldsmith 1999). The problem is that investors cannot agree on how to classify Reuters and therefore how to value the company:

> …even the most seasoned media-sector investors concede that they are less than secure when evaluating Reuters. 'We built a huge [earnings] model…,' said Mark

> Beilby, an analyst in London with Germany's Deutsche Bank. ‘We're confident with it’ for many media companies, ‘but we don’t know whether it works or not for Reuters.’ Other firms express similar uncertainty (Goldsmith, 1999).

The result has been a high volatility on Reuters' stock prices, often with huge losses occurring for no particular reason except the classification uncertainty.

The structural shape of a system of forms also matters. One useful way to think about this issue involves asking what advantages might accrue to a hierarchically-embedded sub-form as compared to a stand-alone form. Organizational sub-forms might be advantageous because of the following processes: lessened need to justify claims made to outside resource controllers (because of possible use of precedence arguments); possibility of protective action by entities from the super-form (because of solidarity and perceived common interests); ease of recruiting members and other resources (because outsiders associate with established super-form); ease of interaction with outside exchange partners and regulators (because partners can use existing sub-units and routines to interact rather than develop new ones). Note that these processes generate hierarchy because of environmental selection but the individuals behind form-based entities may understand such forces and be capable of some limited strategic placement of forms.

Theoretical Processes Generating Forms: Orienting Ideas

Received theory relies on two general ideas about the processes behind the emergence of an organizational form. First, many theorists stress the role of formal institutions in establishing a new organizational form. In reviewing recent research, for instance, Scott (1995: 147) notes that “new industries require the development of a social and political infrastructure that provides working rules, governance structures, and legitimacy.” The infrastructure referred to includes industry associations, professional associations, regulatory bodies and the like. Among other things, these institutions are viewed as providing social order and reducing uncertainty, although some theorists stress the control or domination aspects of prevailing institutions. Second, other theorists—specifically, those working from an ecological perspective—highlight the role of organizational proliferation in establishing a form. The theory of density-dependent organizational evolution states that as the number of organizations using a particular organizational blueprint increases, the blueprint becomes a legitimated organizational form.

Although both approaches claim ample empirical support from a number of studies, a problematic aspect of much research on this question is that typically only established forms are selected for study.[3] Investigation takes the form of retrospective data collection and analysis in searching for the processes that generated and established the form. Of course, such an approach may be very insightful in understanding what occurred and how it happened but it is limited in terms of its ability to establish causality. The problem is that the "independent variables" are identified only after the outcome is known. Thus, these variables may not have produced the outcome—they could be spurious; they also could very well be necessary but not sufficient causal factors. Given the number of replications of each type of study, the latter interpretation seems more plausible than the former.

The research reported here was designed to gain some leverage on this thorny inference problem. We used the two general theoretical approaches described above as theoretical orienting ideas. That is, in examining disk arrays, we looked for and assembled information on both formal institutions and organizational density. We present this data below and use it to inform the theoretical ideas, much as do other studies. However, because the disk array form is not yet fully established, our study avoids—or at least changes—the dominant retrospective analysis problem for at least two reasons. First, disk array production may not eventually become an organizational form. Second, even if disk array producers become a form, by observing before its establishment we might learn something about the sufficient (as well as the necessary) conditions of processes generating it.

Indeed, we think that the study reveals some interesting things about both types of general theoretical approaches. As explained below, it also leads us to advance a new theoretical argument about the establishment of organizational forms, one that we think merits systematic empirical investigation. Before we get to that, however, it is important to understand what disk drive arrays are and how they developed, and to examine the relevant empirical facts associated with both theoretical approaches. We do so in the next four sections, which provide (in order) an introduction to the technology and an assessment of its organizational form status, a historical narrative of its development, a description of the formal associations found in disk arrays, and an examination of organizational density and its underlying vital processes. After setting up the analysis with these facts, we then in the final section confront the theoretical approaches with them.

DISK DRIVE ARRAYS: TECHNOLOGICAL DESCRIPTION AND ORGANIZATIONAL FORM STATUS

The main technical components of a disk array are: (1) a set of disk drives; (2) configuration of the drives into some kind of interdependent system; (3) the interconnect protocols in the system; (4) the storage controller; and (5) the system cache architecture. The business of disk arrays appears even more complicated because arrays are sold with varying degrees of completeness (Disk/Trend, 1998). A number of companies sell subsystems (complete arrays ready to use), but product groups also include boards (array controllers, power supplies and other components without disk drives), and software (an individual software product providing array functionality). Thus, companies may specialize in boards or software, or else they may provide complete systems. Companies may also be independent providers, or else be captive producers making arrays for their own computer systems.

The market for disk arrays is segmented in a number of ways, and firms differ in the scope of their offerings. Arrays are sold in four identifiable primary markets: the computer mainframe array market (e.g., computer reservation systems), the network/midrange multi-user market (the bulk of the disk array market), the single user market, and the specialized high performance market (e.g., video servers, geophysical exploration data analysis). A disk array can have as few as two disk drives or as many as a couple of hundred; most arrays contain fewer than 100 drives.

This market also cuts another way. Array companies serve particular operating environments such as UNIX, NT, NetWare, OS/2, other proprietary environments, and open environments; they may specialize or serve multiple environments. Increasingly, disk arrays comprise elements in a storage system, which is, in turn, an element in a network of interacting servers, specialized servers, and client systems/users. Among their other functions, disk arrays form one of the legs underpinning the internet; without arrays, the internet could not function effectively.

During 1998, 134 companies offered array subsystems, boards or software at one time or another. However, three firms – IBM, EMC, and Compaq Computer – held almost three-quarters of the total market (Disk/Trend, 1998). Led by IBM and Compaq, captive sales accounted for almost two-thirds of industry revenue. EMC was the largest independent supplier, accounting for more than half of non-captive sales, followed by Data General and Hitachi Data Systems. U.S. firms hold 90% of the market.

How do we know that the organizational form for disk array producers is not yet established? As Ruef (1999) notes, organizational forms emerge over time as a social process and a specific form's actual emergence cannot always be readily demarcated by an exact date or a particular organizational event. Nonetheless, it is often possible to ascertain a form's status at any particular point in time and to observe dramatic relative changes in status across two or more points. For instance, in the early 1980s when brewpubs first appeared in various local areas, operators often went to great lengths in developing promotional materials and service routines intended to explain the form to customers and others. The explanations usually included information on what the form was, how it differed from other "neighbor" forms (large mass production brewing firms) and why it was important. By the late 1990s, these explanations have typically fallen by the wayside: individuals know the form by name and easily recognize instances of it by sight and other perceptual tools. The establishment of the brewpub form was propelled by the fact that there was general agreement about the concept among early entrepreneurs, many of whom maintained contact with each other. This consensus by insiders undoubtedly made it easier to lobby for enabling legislation and to launch collective events and formal associations, all of which served to create awareness and understanding among outsiders.

When examining the disk array market today, we do not find much agreement about the organizational form most appropriate for serving it. It is obvious that the public at large does not hold such a conception--many heavy personal computer users are not even aware of the technology, let alone the producer organizations behind it. Relevant outsiders such as security firms' market analysts seldom focus on a disk array (or similarly termed) industry, preferring instead to stay at the more encompassing level of “data storage” (see Tucker Anthony 1998; Hambrecht & Quist 1998). Moreover, these analysts' reports usually do not contain subgroupings based on the disk array form or organizational type; instead individual companies and their particular technologies or product lines (e.g., video or audio streaming, transaction processing, web caching) are described. Looking at insiders to the market, a similar lack of consensus about appropriate form is evident. Companies tend to refer to their particular technology as though it will be at the center of activity in the future. So, for instance, companies refer to themselves variously as involved in “storage,” “storage subsystems,” “RAID,” “disk arrays,” “network attached storage” and others. As the director of product marketing at Maximum Strategy says, “Companies are starting to go away from saying [disk arrays], and are instead talking about what they offer. We provide high bandwidth” (Electronic Engineering Times, 1996).

Companies are also giving the strong impression that they realize the importance of being associated with an institutionalized way of doing things. For example, there are several different classes of disk array products, each stressing different technologies and aspects of their performance. One group of array manufacturers banded together to form a benchmarking group, the Storage Performance Council. As the marketing manager for

Compaq commented, “We discovered there are a fair number of performance numbers around the industry, all quoting different figures, and most of them were drummed up to make a specific product look good” (TechWeb News, 1998). The COO of Mylex claimed the new council will enable the storage industry “to rely upon a set of standard performance benchmarks that successfully measure the performance of products in their targeted working environments” (Electronic Engineering Times, 7/6/98). Such metrics would also “be critical to establishing baseline and improved performance curves that will be meaningful to our customers,” said the director of storage technology at Unisys (TechWeb News, 1998).

HISTORICAL OVERVIEW OF DISK ARRAYS AND RELATED TECHNOLOGY

Disk array technology has been around for many years, but its first appearance proves hard to pin down. Although demand for storage reliability and performance is not new, most analysts think the disk drive array business is. The technology originates in the idea of redundant, or fail-safe, computing when on-line transaction processing began to emerge in the 1960s. Yet, independent companies were slow to offer fail-safe disk storage products; disk drives were bundled with computer systems for which they were specifically designed. Thus a market for disk drive arrays was slow to develop. Disk array technology has co-evolved, in rough chronological order, with fault-tolerant computing, the market for disk subsystems, RAID (redundant arrays of inexpensive disks; “inexpensive” was later changed to “independent”), and storage networks of which disk arrays are one element. We briefly review each. As a guide to the discussion, we provide Figure 1--a timeline that shows key developments regarding the history of disk arrays.

[Figure 1 here]

Fault-Tolerance

Reliability in computing was a problem from the beginning of the computer industry, and initially the only solution was to buy a second computer. This meant that computer users would guarantee total computer uptime with a minimum of delay by having two complete computer systems; when one system broke down, the other was used as a backup.

Military and space programs were the first users of fault-tolerant computing, as well as an important source of R&D funding. Commercial applications, however, appear to be the main drivers of the technology. Airlines were among the first commercial customers for this primitive kind of fault-tolerant computing. In the mid-1960s, it was common for airlines to have three computers to handle their reservations systems, with three more standing by in case something went wrong.

Smaller customer firms could not afford the large fault-tolerant systems. So, a market for less expensive fail-safe computing emerged in the mid-1970s and early 1980s. Pioneered by Tandem Computers, which shipped its first redundant computer system in 1976 (a "multiminicomputer" system), the market for fault-tolerant online transaction processing grew rapidly. Unlike the systems that preceded it, the Tandem computer could (with a high probability) detect and recover from some classes of internal failures -- mostly hardware, but occasionally software errors, too -- before failures affected the processing, the database, or the end user.

Tandem's success and the rising demand for fail-safe and reliable systems invited entry by others. The leading challenger was Stratus Computer, Inc., founded in early 1980 by a group of former Data General executives. A clutch of other companies followed, and by the mid-1980s a dozen or so vendors offered fault-tolerant systems. Most carved out niches rather than confront Tandem and Stratus directly and adopted diverse approaches to fault-tolerance.

With the exception of Hewlett-Packard, conventional systems manufacturers did not move quickly to offer fault-tolerant architectures. However, analysts speculated at the time that new entrants specializing in fault-tolerant computing would nonetheless find a difficult competitive environment because buyers would be concerned about compatibility with existing DEC and IBM systems, which dominated the minicomputer and mainframe markets.

This turned out to be true. By 1985, with the exception of Tandem and Stratus, all fault-tolerant system suppliers ran into trouble, and most exited the market.

As a concept, fault-tolerance focused on the non-stop reliability of the entire computer system rather than the availability of the stored data *per se*, but most fault-tolerant computer manufacturers also built redundancy into their data storage through a technique called “disk mirroring.” Disk mirroring requires duplicate disk drives: if a failure occurs on one disk, an up-to-date, accessible copy of the data will exist on the mirrored disk. This was an expensive solution to data protection in an online environment (as opposed to backing up data using magnetic tape). Because mirroring required a duplicate set of storage devices for keeping a duplicate file of all data, it doubled the cost of storage.

Mirrored disks were generally found *only* in fault-tolerant systems, and no arrays of mirrored disks were marketed independently from systems. Until the late-1980s, no specialist firms made “fault tolerant” storage devices. Although fault-tolerant system producers had the capability to offer redundancy in their disk subsytems, few survived to offer disk drive arrays. By contrast, several of the major systems manufacturers made the transition: IBM, DEC, NCR, Unisys, and Hewlett-Packard.

Disk Subsystems, Disk Arrays and RAID

IBM appears to be the first company to focus specifically on disk arrays as a means to improve the reliability of storage. In 1978 it received a patent for an improved arrangement for recovering and storing data on multiple storage devices in the event one of the devices became inoperable. The disk mirroring techniques then in use, such as those performed by Tandem, were not suitable for correcting or recreating long records that were in error or unavailable; this meant mirroring was not a practical commercial solution when the amount of data transferred involved substantial time to transfer. IBM’s idea was to subdivide a record into a number of segments and store them on different “failure independent” drives. If one drive failed, the segment on that drive could be reconstructed.

IBM was slow to implement this idea, which germinated for almost a decade. In the meantime, some independent companies began to offer software products and disk subsystems employing disk mirroring for the PC and workstation markets. These products typically employed two disk drives and one host; they also had somewhat different objectives than the fault-tolerant systems then in use for mainframes (Computerworld, 1985). Where fault-tolerant systems emphasized the role that mirrored disks played in the protection of critical data, mirroring was increasingly said to lead to better performance as well. Industry participants and the business press began to refer to bundles of mirrored disk drives as disk subsystems or, increasingly, disk arrays.

The first reference to disk arrays (that we can find) appeared in 1986 and was a software solution offered by Twincom, a Dutch-based company. It was followed by seven other disk array offerings in 1987, including those of six non-captive companies: 1776, Atlantic Microsystems, Core International, Ford/Higgins, Maximum Strategy, and Micropolis, a disk drive manufacturer.[4] Yet, most observers date the "origin" of the disk array "industry" to a December 1987 U.C. Berkeley technical paper (Patterson, Gibson, and Katz, 1987), which helped companies view and classify disk array technology in a common way. The Berkeley research was sponsored by the National Science Foundation and supported by IBM, Sun Microsystems, Intel and DEC. Its point of departure was clearly IBM's 1978 patent.

The Berkeley paper addressed a growing problem in data storage: storage devices could not keep up with the speed of faster and faster CPUs; disk drives were a major input/output (I/O) bottleneck in achieving rated processor speed. Some disk drive producers argued that lashing together several small hard disk drives electronically in a single subsystem would be the best solution to the problem. More specifically, performance would be enhanced and costs reduced by combining multiple inexpensive 5.25-inch disk drives in lieu of the higher capacity 8-inch and 14-inch drives that were used in mainframe and fault-tolerant computer systems.

The Berkeley team argued that the only way to increase performance to the required levels was by writing data across several disks ("striping") rather than to one. Proprietary software would let the system regenerate information stored on a drive that failed. Striping was an attractive alternative to mirroring because it increased overall throughput: data could be written faster than would be possible if one drive were dedicated to duplicating data. Also, potential customers, in principle, liked the idea of smaller disks because they were cheaper, generated less heat and used less floor space than the standard mainframe disk size of 14 inches.

The Berkeley solution was a set of five arrangements of arrays of small disks offering various levels of protection and known under the generic name RAID.[5] These are described in some detail in Appendix A.

Despite these developments, disk drive makers and system builders still moved cautiously with disk array and RAID technology. A common comment among experts was, "It's really hard to say whether or not this is a market" (Electronic Business, 1988). Many thought disk arrays only represented an opportunity for companies already in the systems or disk drive business; a company that based an entire business plan on it was thought to have a lower chance for survival. As the founder of Auspex said, "How do you get in between [disk] drive and systems companies and make money?" (Electronic Business, 1988). Industry experts argued that widespread acceptance of disk arrays would have to contend with small profit margins, an education process to get the market excited (distributors and value-added resellers did not know what disk arrays even were), as well as the threat posed by the established computer companies. A more fundamental concern was whether the technology represented a new product opportunity or only the repackaging of existing disk drives with little added value. As a Dataquest analyst commented at the time, "The sales effort to educate the prospective customer base regarding the benefits of array processing will be expensive" (Electronic Business, 1988).

The RAID disk systems that initially appeared after the publication of the Berkeley paper served the small server and minicomputer markets (Computer Weekly, 1995). But many analysts and mainframe computer users thought these initial entrants were ahead of the market by about two years because of concern about the reliability of 5.25-inch drives and the fact that they could not tolerate the high data rates required by mainframe users.

In fact, the market for disk arrays did not begin to attract many entrants until the early 1990s. In August, 1990, EMC Corp. became the first company to ship a RAID product into the mainframe market (Electronic News, 1990; Computerworld, 1990). After that fast start, EMC became, and remains, the largest independent producer of disk arrays, with 21% of the worldwide market in 1998. Follow-on entrants targeted a variety of users: mainframe, workstation and microcomputers, especially for critical database applications on PC-based network servers. But minicomputers and multiuser systems comprised the largest markets for arrays; Compaq, Lomas Data, Storage Concepts and Sequoia were early entrants that served this market. Except for EMC, only Hitachi, Control Data Systems, Unisys, and Storage Technology offered RAID for mainframe storage before 1994.

By the mid-1990s, RAID systems were finding market appeal beyond their initial application for the stock exchange, airlines, medicine and major retail. There are several reasons for this. First, the disk drive industry had improved the reliability and performance of small form factor disk drives, and had achieved tremendous economies of scale in producing them. These developments translated into higher performance disk arrays using smaller and cheaper components. Second, operating systems and software applications consumed ever-increasing amounts of storage, driving demand for capacious systems. Finally, companies increasingly demanded continuous online access to data, the so-called 24x7 phenomenon.

Network Attached Storage, Storage Area Networks

During the last five years, disk arrays employing RAID technology have emerged as central elements in corporate electronic communication networks. Companies are increasingly integrating data storage with distributed computing systems in an effort to access data from anywhere in a network at any time. This requires RAID systems that enable fast data transfer rates while maintaining continuous fail-safe access throughout the network.

Disk array companies have come up with two approaches for incorporating storage into these networks: Storage Area Networks and Network Attached Storage. Disk array systems were originally server- or host-attached storage devices – components in so-called Storage Area Networks, or SANs. A SAN involves a separate, high-speed network to move information among storage devices. In a SAN, users connect to the network through a server and go through the server to access data. By contrast, Network Attached Storage, or NAS, refers to a disk array connected directly to the network backbone via some generally accepted cable technology such as ethernet, Fibre Channel, or asynchronous transfer mode. NAS enables users to access data more quickly without affecting the server's processing function.

As these two approaches exemplify, technological competition and uncertainty persist. Despite the establishment of RAID levels, a decade of product offerings, and the rapidly growing demand for disk arrays by companies that want to integrate their Internet data with traditional business data, analysts still describe the RAID market as fragmented and confusing to potential customers.

INDUSTRY ASSOCIATIONS RELATED TO DISK ARRAYS

We now turn to industry associations, pursuing evidence relevant to the suggestion of some scholars that collective action fosters the emergence of organizational forms (Aldrich and Fiol, 1994; Reuf, 1999). Specifically, industry associations might potentially help fashion and sustain new organizational form identities. In the disk array market, two major industry groups have emerged: the RAID Advisory Board (RAB) and the Storage Networking Industry Association (SNIA). Table 2 shows the membership in RAB and SNIA. Disk array companies show considerable variation in membership, with the two groups reflecting overlapping rather than nested sets of endorsed organizations, as we describe next.

(Table 2 about here)

RAID Advisory Board

In August 1992, the RAID Advisory Board was formed with ten members.[6] Their main motivation was to reduce the confusion caused by disk array companies, which were making often contradictory or inconsistent claims about the technology. RAB's objectives were to clarify RAID terminology, educate the market (particularly the engineering community), and promote common interests among operating system developers and standards bodies for the fledgling industry.

One of the most pressing problems facing the Board was customer confusion regarding RAID levels and performance benchmarks. The five Berkeley RAID levels had not been officially sanctioned by any standards groups, and RAID levels were being implemented in different ways. It was not clear, for instance, that Level 1 was the same across products; a RAB-certified RAID Level 1 product, by contrast, would inform the user exactly what the product does. RAB produced a 100-page book of RAID definitions and sold more than 1,000 copies during the first three months after its release.

Nonetheless, confusion about RAID levels continued. A particular misunderstanding was that RAID levels were hierarchical, that the higher the RAID level, the better the system. Said RAB's chairman, "[T]he RAID levels have done more to confuse people than anything else" (Midrange Systems, 1993). In spite of 70% growth in disk array sales in 1994, he claimed that the "increasing activity of RAID technology is -- not surprisingly -- resulting in confusion and misperception" (Computer Design, 1994).

Another cause of confusion was the host of new products introduced claiming *additional* RAID levels. "Bogus levels have been generated to describe products," warned the chairman of the RAB, in 1993 (Corporate Computing, 1993). Just when RAB was trying to get people to understand the most prevalent levels (1 -- disk mirroring, 3 -- striping, and 5 --

parity and striping), new levels (6, 7, 10, and 51) further cluttered the market. These levels were proprietary solutions not recognized by RAB. Demystifying RAID technology and getting agreement on some standards were thought the key to increasing RAID sales. According to RAB's chairman, "Standardization is probably the biggest issue we face. It's critical to RAID's growth, and it's an exhaustive process" (LAN Times, 1995). But an analyst at Dataquest Inc. said that the RAB faced a tough task in trying to convince the 100 or so RAID vendors to settle on standards. She said, "We'll probably see de facto standards, based on who can build a better mousetrap" (LAN Times, 1995).

RAB had some success in raising user awareness and understanding of disk drive arrays and RAID. "There's no one out there like them [RAB]," said an analyst for International Data Corp. "The industry does need some sort of consortium to set standards for vendors, and the RAB provides an ample forum or discussion of the topic among manufacturers and users" (LAN Times, 1995). But the role of storage in computing continued to change rapidly, and the disk array market was not settling on RAID as a central identity. In 1994, 55 companies were members of RAB; by 1998 approximately 40 were, less than a third of the population of disk array makers.

Storage Networking Industry Association

As networked computing and the Internet gained in importance, disk array makers and industry analysts began to focus more on the role of data storage within a network rather than storage *per se*. Disk array producers, who had been struggling to create a clear identity, suddenly found themselves reshaping their images. They still offered disk arrays employing RAID technologies, but they began to be evaluated in terms of their role in broader networks.

The confusion with these approaches, combined with the expectation of enormous sales growth, sparked the formation of the Storage Networking Industry Association (SNIA). Formed in October 1997, the SNIA is more broadly focused than RAB, and its 80 members come from the wider IT community; it sees itself as leading the way for the storage networking "industry."

The SNIA has two principal objectives. One goal is to bring together into a single community the various information systems firms, systems integrators, and vendors to focus on education and training in networked storage systems. A second objective is to facilitate standards for communication between servers and storage devices. As the SNIA chairman summarized, "What we're looking at doing is helping to provide the infrastructure, including the specifications and the first standard, for tying together storage systems on the network, being able to aggregate storage and present it to users from a showable perspective on the network. We're expecting to provide benefits across the whole industry, [including] providing education and an infrastructure for the industry to support this new paradigm of shared-network storage systems." (Computer Dealer News, 1998).

Although technical challenges abound, collective action is also required from a marketing perspective. According to the SNIA president, "Storage community leaders have not figured out how to develop a market…We now must recognize that networking and storage are systems industries. We need to effectively market our products by establishing connections with the IT community….SNIA is attempting to build an IT organization, not merely an ad hoc standards organization" (Mass Storage News, 1998).

Both RAB and SNIA have succeeded in generating awareness of disk arrays, but disk array producers still have a rather diffuse identity. In some ways, technological change has outpaced RAB's efforts to promote the RAID concept; the rapid emergence of distributed computing has made disk arrays only one element in communication networks. For its part, the SNIA focuses on a broader organizational community than disk arrays, but the association was formed too recently to have had much visible impact. Either way, collective action has so far failed to result in a clearly identifiable organizational form.

ORGANIZATIONAL DYNAMICS IN THE DISK ARRAY MARKET

Now we examine organizational activity in disk array production, using systematic data collected on firms participating in the disk array market. Our data collection effort attempted to identify every firm that ever sold a disk array product and the dates when it entered and exited the market (if it did so by the end of our observation period, year's end 1998). The procedures and sources we used are described in Appendix B. Because of imprecision in the sources, the data we use are precise only to the year and not to months or specific dates within years.

Our attempt to reconstruct the organizational history of the industry yielded a total of 257 firms. This count begins in 1986 with the first firm in the market, Twincom, which entered with software for disk mirroring.[7] It covers all firms known to offer disk arrays up to and including the eleven new entrants in 1998: Acard Technology, ADI, Chaparral Network Storage, Creative Design Solutions, Engrows, Lexias, OneofUs, Sagitta, SMS Data Products Group, SoftRAID, and Synapsys Digital.

Figure 2 shows a plot of annual organizational density of disk array producers across the history of the technology. It includes all firms that entered the market at any time, regardless of level of activity or other activities of the firm. Firms are removed from the density count whenever they exit the market, regardless of the reason.

(Figure 2 about here)

The temporal trajectory of density visible in Figure 2 resembles that of many coherent organizational populations. As Carroll and Hannan (2000) review, the density pattern of initial slow growth, followed by rapid growth and then stabilization or decline appears in very diverse sets of organizations operating in widely different contexts. What differs across organizational populations is the timing of the various phases of the trajectory and their peak levels. Many of the populations reviewed by Carroll and Hannan (2000) exhibit much greater prolongation in the different phases than that seen here. The collection of disk array producers is notable for the historical speed with which the "full" pattern appears. Perhaps this is a result of the fast pace of technological change in this market. Perhaps too the recent downturn is only a temporary phase that will eventually lead to renewed fast growth in density.

What do the organizational dynamics of entry and exit look like? Figure 3 plots the annual number of firms entering the market for disk arrays while Figure 4 shows the number of exits per year. Entries into a market often show great fluctuations from year to year (see Carroll and Hannan 2000). By contrast, the plot of entries of array producers shows a more pronounced series-long pattern than is typical of many familiar organizational populations. Nonetheless, although the years 1992-94 generated more entries than those periods before or after, the series is short enough to warrant caution in making inferences. The plot of annual exits looks more typical; it suggests that many of the bulge entrants in 1992-94 operated for 2-3 years before exiting. However, it would be premature to draw conclusions about tenure dependence in the market from these aggregate data.

(Figures 3 and 4 about here)

So, at this first casual glance, the disk array market looks similar to many established organizational populations in terms of density and exit (Carroll and Hannan 2000). The entry process looks like perhaps it might be somewhat different. Examining entry in greater depth only reinforces this perception. In Table 3, we show a tabulation by entry mode for those entrants to the market for which we could find such information. The table also shows counts by origin industry for those firms not starting anew in the disk array market.

(Table 3 about here)

In perusing the table, two features of the disk array population stand out compared to most others described in the ecological literature. First, the number of *de alio* (lateral) to *de novo* entrants is extremely high. Our knowledge of many of these firms suggests strongly that most of them continued with their prior activities for an extended period after they entered the disk array segment. Second, among the *de alio* entrants, the diversity of origin industries is very high, ranging from highly technical firms in computers and motherboards to other fairly non-technological firms involved in distribution and retail. Both factors figure into the theoretical interpretation discussed below.

RECONSIDERING THE THEORETICAL ISSUES: LIMITATIONS AND PROPOSALS

As discussed at the outset, our purpose in this paper is to examine how and when organizational forms emerge. Following Polos et al. (1999), we define an organizational form as a recognizable code that possesses rule-like standing. More precisely, a form is an external identity code. This means that when an applicable organization is perceived as violating the code, outsiders discontinuously devalue the organization in their assessments and impose sanctions. Identity codes for organizational forms typically consist of abstract features as well as composition rules about appropriate combinations of particular features.

Contemporary organizational theory envisions two very different kinds of processes behind the emergence of an organizational form. Institutional theorists emphasize how formal institutions, such as industry associations, professional associations, and regulatory bodies, assist in the establishment of a new organizational form. Theorists view these institutions as providing social order and reducing uncertainty, as well as control or domination. Ecological theorists focus on how proliferation aids in establishing a form. The theory of density-dependent organizational evolution holds that as the number of organizations using a particular identity code increases beyond a critical minimal level, the code becomes an organizational form (Polos et al., 1999.) In applying the theory, ecologists count the number of organizations (density) in a market across time.

Our historical review of the market for disk arrays shows the validity of both theories in describing some of what occurs in the early stages of a market or industry. Disk array technology develops out of murky origins and spawns a large international market. Along the way, two large and important formal associations crop up and develop. Much as institutional theory would predict, these associations concern themselves with problems of order, especially standardization. Similarly, as ecological theory would predict, the number of producers entering the market rises steadily over the early years, eventually slowing down and stabilizing and, finally, declining slightly. That is, the characteristic organizational density trajectory appears, albeit on a short timeline.

From the point of view of the theories, what is problematic in the case is not what happened in this market but instead what did not happen. Specifically, there is little consensus on what it means to be a disk array producer, there is no taken-for-granted way to organize for this market, and there is no established identity code used by outsiders to evaluate whether a firm qualifies as a legitimate disk array producer. In sum, there is no organizational form for disk array production despite the appearance of formal institutions and the expected trajectory of producer density. So, the theories fall short in accounting for this case.

It is important to recognize that this shortcoming does not mean that the theories are unhelpful. The factors pinpointed by each can still be viewed as necessary conditions for the establishment of an organizational form but they no longer seem highly plausible as sufficient conditions. This observation leads naturally to questions about whether the case can provide insight for further theory development in this direction.

In examining the possibilities for theoretical elaboration, a number of routes seem potentially productive. First, the theories could require refinement in terms of the timing of either or both processes and eventual form generation. That is, a delay in the timing of purported effects might need to be more fully specified. Unfortunately, the case provides few hints about this issue, which seems more easily addressed in comparative research. Second, there could be a technological component to the problem. For instance, it could be that a greater consensus about the underlying technology is required in addition to the formal institutions and rise in organizational density. Although attractive to some, we are dubious of this kind of explanation because it seems easy to point to many established organizational forms lacking such a common technological foundation (e.g., the computer manufacturing form or the early automobile manufacturing form). Third, it could be that institutional theory about formal institutions and associations needs further qualification. One way to refine this theory might be to relate organizational form generation to the level of fragmentation (both formally and informally) in the relevant institutional environment (Scott and Meyer, 1989). For example, because there are two competing formal associations and there is no solid consensus within them or across them, the disk array market has a somewhat fragmented institutional environment. The difficulty with this approach is that it is difficult to distinguish clearly between institutional fragmentation (the independent variable) and the lack of consensus about form (the dependent variable). In reality, the two often coincide very closely.

What to do? We think the disk array case suggests an alternative fourth theoretical avenue. In our view, the observed high diversity of origins and other activities of those organizations operating in this market are highly pertinent. More precisely, we suggest that the disk array producer organizational form has not taken hold because disk array producers come from a heterogeneous set of origin industries and often retain operations in those industries, perhaps still deriving the bulk of their revenue therein. By this view, the problem resides in the identity basis of firms: so long as firms in the disk array market derive their primary identities from other activities and so long as there are few highly focused firms deriving their primary identity from disk arrays, then the disk array producer identity seems unlikely to cohere into a code or form of its own. We imagine that there are many other cases of established markets without specific organizational forms because the supplier firms are primarily engaged in other identity-defining activities, which could be vertically or horizontally related to the focal market.

Consider also a counterfactual example from another context, the beer market. Suppose Anheuser Busch, Miller, Coors and others had been the first and only corporations to open brewpubs.[8] If they had done so, then we suspect that a market for freshly brewed beer would have still developed and perhaps even flourished more than it did--but a new organizational form would not have. Instead, there would likely be many local beer outlets operating as appendages of the large industrial brewing firms and without much of a separate identity. The identities of the major brewing companies would dominate this market as they do the packaged beer market. In other words, an organizational form for brewpubs may not have emerged without the many focused entrepreneurial firms doing only this activity. And note that in this counterfactual scenario, the major beer producers look fairly similar on many dimensions; they do not possess the heterogeneity of disk array producers.

If these speculations are valid, then they lead to a reformulated specification of the density-dependent process thought to generate an institutionalized organizational form. More precisely, we propose that a legitimated organizational form emanates from the density of either focused or *de novo* producers in a market rather than total density. The core idea is that identity of a form derives from the aggregated identities of individual organizations. When the identities of individual organizations are all similar and focused on a particular activity, then they will gel more readily into a collective identity. When organizational identities are diverse and diffuse, then emergence of a collective identity is more problematic.

An advantage of this formulation is that it can be readily incorporated into extant models and theories of density-dependent legitimation. This formulation is also more consistent with the empirical facts of the case. Consider, for instance, Figures 5 and 6, which show respectively the annual density and entries of *de novo* producers into the disk array market. Unlike total density, *de novo* density does not rise to a stable point and then subside. Rather, *de novo* density appears to be still in a growth phase. More importantly, because it does not level off and is still rising upward, the trajectory of *de novo* density does not give the impression that the organizational form should be legitimated; it suggests instead that the organizational form is undergoing institutionalization and may not yet be legitimated. This suggestion is, of course, consistent with what we observe in the market. So, although further empirical research on the issue is still needed, we think the case warrants our theoretical conjecture about reformulating the density specification.

(Figures 5 and 6 about here)

Although this formulation appears theoretically sound and empirically consistent with the facts of the disk array case, it should be noted that it contradicts exactly one drawn from another popular perspective on legitimation. Precisely, the so-called socio-political view of legitimation holds that endorsement by powerful actors yields advantages to organizational forms and aids in the process of legitimation (Scott 1995). It follows logically then that if and when larger established (powerful) organizations enter a market, then legitimation should be enhanced. IBM's entry into the PC market is a well-known case that seems consistent with this argument. In terms of organizational density by entry mode, the prediction most consistent with this view would be that *de alio* density contributes the greatest to legitimation. This is because *de alio* entrants will usually be larger and more powerful than *de novo* entrants.

We also think that the formulation offered here might require qualification with respect to spatial agglomeration. Producers of disk arrays are geographically scattered around the U.S. Surely, such spatial dispersion makes it more difficult for a collective identity to form and hence may make homogeneity of individual producer identities more important. Looked at from the other side, spatial agglomeration may facilitate collective identity construction. Specifically, agglomeration may make it possible for diverse actors to recognize and act on their commonalities; it may also make it easier for outsiders to see and identify the form. An example that appears to accord with this speculation is New York City's emerging Silicon Alley. However, only additional systematic study can tell for sure the value of this possible qualification.

CONCLUSION

Most analyses of strategy and organization presume a well-defined market or industry and then analyze the various locations of firms (and their products) within the industry's structure. This is the approach taken, for instance, in Michael Porter's (1980) influential framework for the competitive analysis of industry. It is also the presumption behind much antitrust regulation.

Yet, in most industries, the boundaries for competition change as technologies evolve, standards become institutionalized, regulations get developed and competitors make bold moves to expand their domains. Such industry re-definition often occurs with great speed in modern industries but it can also be seen in traditional industries if a long enough historical perspective is used. In both cases, it is interesting to know that contemporaries often fail to grasp the changes about to overcome them. Sometimes developments thought to be important prove to be trivial; at other times, things thought to be minor turn out to have revolutionary impact. In other words, when viewed prospectively industry evolution appears to show great uncertainty; it is mainly in retrospect that the direction of industrial change appears orderly and rational.

The aspect of industry evolution that motivated the research reported here concerns organizational forms. The project aims to address a basic question related to the evolution of organizational forms and their constituent firms: When and where will a new organizational form emerge? Contemporary organization theory suggests two primary processes. The first is that formal institutions such as industry associations and standard-setting bodies will lead to a taken-for-granted organizational form. The second is that rising organizational density (number of organizations in a population) will generate a legitimated organizational form.

Our historical study of the disk array market and its associated technologies reveals the limitations of each of these theoretical arguments. Although we found significant collective activity in association-building and standard-setting among disk array producers, these did not readily lead to an organizational form. Likewise, the widely observed trajectory of a rapid increase in organizational density followed by stabilization did not generate an organizational form. In sum, both processes have operated as expected at the start of the population but neither has yet produced the expected outcome of an organizational form.

Why this is the case must, of course, be considered a speculation. But in our view, a major part of the answer lies in the diversity of origins and other activities of those organizations operating in this market. As long as disk array producers come from a heterogeneous set of origin industries and continue to operate in those industries (deriving the bulk of their revenue from such), then it is hard to imagine a disk array producer organizational form taking hold. Why? The problem is one of identity: if firms in the market derive their primary identities from other activities and there are few highly focused firms deriving their primary identity from disk arrays, then the disk array producer identity cannot cohere into a code or form. A similar type of "market without new organizational form" might have developed for local freshly brewed beer if Anheuser Busch, Miller, Coors and others had been the first and only corporations to open brewpubs. (And this counterfactual scenario uses producers who collectively do not possess the heterogeneity of disk array producers.) That is, an organizational form for brewpubs may not have emerged without the many focused entrepreneurial firms doing only this activity. Instead, we might have had many local beer outlets operating as appendages of the large industrial brewing firms and without much of a separate identity.

If our speculations from this case study are valid, then they suggest a reformulated specification of the density-dependent process thought to generate an institutionalized organizational form. More precisely, we propose that a legitimated organizational from emanates from the density of either focused or *de novo* producers in a market rather than total density. An advantage of this formulation is that it can be readily incorporated into extant models and theories of density-dependent legitimation. We also believe that this formulation is more consistent with the facts of the case we studied but, of course, only systematic empirical research will determine its ultimate value.

As this study shows, the investigation of how an organizational form emerges and evolves requires the collection of data that goes back to its origins. It is also important to study more than one industry so that general processes can be identified. A larger project, of which this study is a part, was designed with these concerns in mind. We plan to study and collect data on each of the major technologically-defined segments of electronic data storage. These include: floppy disk drive manufacturing; hard disk drive manufacturing; disk array systems, optical storage manufacturing, and tape drive manufacturing. These segments are similar enough in function to allow meaningful comparison; they also develop in distinct enough ways to warrant individual study. The segments are at various stages of development: floppy drives are a mature segment facing diminishing demand; tapes are also mature but still undergoing innovation; hard disk drive manufacturers compete primarily on the basis of cost but must nonetheless keep up with unparalleled rates of technological change; and the disk array business, perhaps the most interesting of all from an evolutionary viewpoint, is in great flux with competitors entering from bases of both of hardware and software expertise. A major additional feature of this design is the potential substitutability of various products across the segments: this will allow assessment of the ways that competition processes affect the boundaries of industries and organizational forms. The project will proceed through both structured and unstructured research activities examining each of the primary segments.

REFERENCES

Aldrich, Howard E. and Marlene C. Fiol. 1994. “Fools rush in? The institutional context of industry creation.” Academy of Management Review 19: 645-670.

Business Wire. 1990. “Two years ahead of its competitors' projected delivery dates, EMC unveils first 'RAID' computer storage system.” September 25.

Carroll, Glenn R. and Michael T. Hannan. 2000. The Demography of Corporations and Industries. Princeton: Princeton University Press.

The Christian Science Monitor. 1990. “‘Electronic brain’ worthy of name. February 21.

Computer Dealer News. 1998. “Storage firms troll for systems integrators; storage systems for computer networks.” February 9.

Computer Design. 1994. “Look out! RAID! Redundant Arrays of Inexpensive Disks.” Jan.

Computer Weekly. 1995. “Storing up trouble; IBM's problems in the disc array market.” May 11.

Computerworld. 1984. “Nohalt announces 16-bit Reliant system.” May 21.

Computerworld. 1985. “DG boast high availability; Eclipse supermini features fault tolerant-like methods.” November 25.

Computerworld. 1987. “Thinking Machines super joins high-speed elite.” May 4.

Computerworld. 1990. “EMC plans compact disc for mainframes.” September 24.

Corporate Computing. 1993. “Right as RAID.” May.

Douglas, Mary. 1986. How Institutions Think. Syracuse: Syracuse University Press.

Electronic Business. 1988. “Disk arrays: here come the disk drive sextuplets; electronically bundled disk drives.” November 1.

Electronic Business. 1993. “Upstart RAID makers thrive against major drive OEMs.” Aug.

Electronic Engineering Times. 1996. “RAID-certification program completed.” January 2.

Electronic Engineering Times. 1996. “Drive vendors downplaying RAID 'levels'.” Mar. 25.

Electronic Engineering Times. 1998. “First one will focus on disk subsystems – Storage industry collaborating on benchmarks.” July 6.

Electronic News. 1988. “14-inch DASD sales climbing.” May 16.

Electronic News. 1990. “Disk storage unit introduced by EMC.” September 10.

Goldsmith, Charles. 1999. “Investors are split over Reuters identity.” Wall Street Journal, December 20.

Hambrecht and Quist. 1998. Data Storage Earnings Preview for Q1 1998. Technology report by Todd Bakar and Patrick Ho. San Francisco: Hambrecht and Quist.

Hannan, Michael T., and Glenn R. Carroll. 1992. Dynamics of Organizational Populations. New York: Oxford University Press.

http://www.snia.org/news/press010898.html 1998. “IS, industry leaders form new trade association.” January 8.

InfoWorld Daily News. 1999. “Sun promotes storage management standard.” July 16.

LAN Times. 1992. “RAID Advisory Board pushes products, not promises.” September 14.

LAN Times. 1994. “RAID vendors hope to end confusion, spark interest.” July 25.

LAN Times. 1995. “HP's AutoRAID takes complexity out of RAID.” June 5.

Mass Storage News. 1998. “Storage Networking Industry Association sets agenda, announces September conference.” July 16.

Midrange Systems. 1993. “Lifting the veil; Advisory Board wants to standardize redundant array of inexpensive disk technology.” September 28.

Network Computing. 1997. “Safeguarding Network Data: This Is A RAID.” June 15.

Network Computing. 1994. “Storage--yet another logo.” October 1.

Optical Memory News. 1998. “Disk drive prices dictate RAID revenue.” February 24.

Patterson, David A., Garth Gibson and Randy H. Katz. 1987. "A case for redundant arrays of inexpensive disks," Report No. UCB/CSD 87/391, December, Department of Electrical Engineering and Computer Sciences, U.C. Berkeley.

PC Week. 1992. “RAID technology gathers steam: interest in arrays grows as PC LANs, storage concerns multiply.” December 7.

Polos, Laszlo, Michael T. Hannan, Gabor Peli and Glenn R. Carroll. 1998a. "Form and identity: On the structure of organizational forms," Presented at the 14th Colloquium of European Group on Organizational Studies, Maastricht, the Netherlands.

Polos, Laszlo, Michael T. Hannan, Gabor Peli and Glenn R. Carroll. 1998b. "On the evolution of organizational forms: Formal models," Presented at the annual meetings of the American Sociological Association, San Francisco.

Polos, Laszlo, Michael T. Hannan and Glenn R. Carroll. 1999. "Identities, form and populations," Unpublished manuscript, Stanford University.

Porter, Michael E. 1980. Competitive Strategy. New York: Free Press.

Ruef, Martin. 1999a. "Social ontology and the dynamics of organizational forms: Creating market actors in the healthcare field, 1966-94," Social Forces 77: 1405-1434.

Ruef, Martin. 1999b. "The emergence of organizational forms: A community ecology approach," Unpublished manuscript, University of North Carolina.

Scott, W. Richard. 1995. Institutions and Organizations. Thousand Oaks CA: Sage.

Scott, W. Richard and John W. Meyer. 1983. "The organization of societal sectors," Pp. 129-54 in J.W. Meyer and W. R. Scott (eds.), Organizational Environments: Ritual and Rationality. Beverly Hills: Sage.

Swaminathan, Anand and James B. Wade. 1999. "Social movement theory and the evolution of new organizational forms," In K. Schoonhoven and E. Romanelli (eds.), The Entrepreneurship Dynamic in Industry Evolution. Stanford CA: Stanford Univ. Press.

TechWeb News. 1998. “Storage Industry Collaborates On Benchmarks.” July 2.

Tucker Anthony. 1998. The Data Storage Market. Equity research report by Maak Kelleher and Phil Muscatello. Boston: Tucker Anthony.

VARBusiness. 1996. “Know your RAID? You've got it made -- The explosion in network storage needs is driving business to VARs who understand RAID technology.” Jan. 1.

VARBusiness. 1998. “LAN giant does storage -- 3Com senses strong appeal in Storage Area Network market.” November 9.

APPENDIX A: DESCRIPTION OF RAID TECHNOLOGY

RAID technology is defined according to seven levels. Each level specifies a different disk array configuration and data protection method, and each provides a different level of reliability and performance. In reality, only a few of these configurations are practical for most online transaction processing systems, file servers, and workstations. These seven levels are briefly described below:

- Level 0. Nearly all disk arrays include this mode. RAID 0 distributes the data across all the disks in the disk array configuration. Since there is no redundancy, the capacity utilization is 100 percent.
- Level 1. Write data is “mirrored” on two separate disk systems. This means that every byte of storage used on a system requires that an equal amount of storage must be provided as a mirror.
- Level 2. This level of RAID has become obsolete. It is seldom mentioned in most descriptions of RAID technology, and most vendors no longer sell it. RAID 2 provides a greater than necessary level of redundancy.
- Level 3. RAID 3 is often referred to as a parallel array because of the parallel method that the array controller uses in reading and writing to the disk drives. For RAID 3 it is necessary to provide two or more data disks plus an error correcting code disk. Data is dispersed or “striped” across at least two of the data disks. Each drive is connected to a dedicated SCSI channel, which further ensures the performance. If a failure occurs, the data is reconstructed based on the parity data on the other disks. The advantage of this scheme is that, compared to RAID 1, redundancy is achieved at a lower cost.
- Level 4. With RAID 4, data are block striped, not bit striped as with RAID 3. Blocks of some arbitrary size are recorded on each data disk. Parity values are calculated and stored on the single check disk in the same way they are for RAID 3. The cost of redundancy is the same as RAID 3 since an extra disk must be used to store parity information. Because there is one parity disk, writes are slow since only one write can occur at a time. Every

write has to write to the same disk. Block striping makes for faster reads since any request for information smaller than the size of a block can be satisfied with a read from one disk. This makes it possible to respond to small reads quickly, and also to respond quickly to multiple concurrent reads. Thus, block striped RAID arrays are effective in transaction processing environments where small pieces of information predominate.

- Level 5. Like RAID 3, RAID 5 distributes the data blocks over the disk drives in the system, but unlike RAID 3, the error correcting code data is also distributed across all the drives. With this configuration reads and writes can be performed in parallel.
- Level 6. RAID 6 not only mirrors data (like RAID 1) but also stripes the information. Because the data is striped, the performance is very similar to that measured in a RAID 0 configuration. The penalty for use of this configuration is that 100 percent more disk space is required.

Note that "level" does not mean there is a strict hierarchy, with level 5 always better than level 1. The different levels are just different techniques for data protection with different cost and performance tradeoffs. Customers evaluate each level of RAID individually. For example, once firms reach a certain threshold of storage capacity, it is unusual for them to mirror (RAID 1). If firms are rapidly expanding their storage needs and adding a large number of disk drives, they might use RAID 5. But RAID 5 "stripes" data one block at a time across all the disks, which can cause certain applications to suffer a "write penalty" and slow system performance. Different RAID levels serve different purposes, with different advantages and disadvantages.

APPENDIX B: DATA SOURCES AND CODING PROCEDURE FOR ORGANIZATIONAL POPULATION DATA

The data we use come primarily from Disk/Trend, Inc., a market research company in Mountain View, California. Disk/Trend publishes annual reports on disk drive arrays, as well as other kinds of storage. The first Disk/Trend report on arrays was published in 1993. The reports cover every company that makes complete subsystems, boards, or software specifically intended to permit disk drives to operate as an array. The reports publish firm-level data on revenues and unit shipments for the largest firms in the industry and in a specific market segment; typically firm-level data cover 90% or more of revenues and unit shipments but only 20% of the companies. The reports also list specifications for each product a company ships, and the date of its first shipment.

In addition, we compiled event histories for each company identified by Disk/Trend as an array manufacturer. These histories were generated by extensive library and online searches, which also turned up a few companies that made disk drive arrays prior to the publication of Disk/Trend. In some cases, the event histories revealed shipment dates that preceded those listed in Disk/Trend and provided more accurate dates for exit from the array market.

Entry dates were coded as the date of first product shipment. Exit dates were coded as the last year a firm was listed in Disk/Trend or else the last year it made a disk drive array if we had more detailed information. We also determined whether a firm was a *de novo* or *de alio* producer. If *de alio*, we also coded the industry the firm was in at the time it shipped its first disk drive array. When firms made more than one product, we classified it according to its dominant market (e.g., computers).

Table 1. Worldwide Sales Revenues of Disk Drive Arrays (in millions of US dollars)*

	1992	1993	1994	1995	1996	1997	1998
Single user systems	2.8	16.7	16.0	32.5	38.6	32.7	38.4
Networks/ Midrange systems	1207.5	2506.5	3821.0	5418.9	6704.3	7884.1	9210.4
Mainframe systems	259.8	867.4	1853.0	4263.3	3958.4	3556.8	3221.0
Specialized high performance systems	17.9	44.8	54.6	74.1	58.3	97.2	106.8
Total	1,488.0	3,435.4	5,744.6	9,788.8	10,759.6	11,530.8	12,576.6

* Sources: DISK/TREND Reports for 1993-1999.

Table 2. Disk Array Companies by Membership in Industry Associations, 1998

Company Name	**RAB**	**SNIA**
Adaptec,Inc	X	
ADIC		X
Amdahl Corporation	X	X
American Megatrends,Inc	X	
Ancor Communications		X
Andataco	X	
Ark Research Corporation		XA
Artecon, Inc.		XA
Auspex Systems		XA
Axis Communications		X
BellSouth		XA
BMC Software		X
Boole & Babbage		XA
Box Hill Systems		XA
Bull		XA
Carnegie Mellon University	X	
Chaparral Technologies		X
CLARiiON, A Data General Company	X	
CMD Technology, Inc.	X	X
Comm Vault Systems		X
Compaq Computer Corporation		X
Computer Associates International		X
Computer Network Technology		X
Computer Technology Review	X	
ConvergeNet Technologies		X
Cope AG		XA
Creative Design Solutions		X
Crossroads Systems		X
Data General Corporation	X	X
Digital Equipment Corporation	X	
Disc,Inc		XA
Disk/Trend,Inc	X	XA
ECCS,Inc.	X	
EMC Corporation	X	X
ENDL Consulting	X	
Eurologic Systems Ltd.	X	
Exabyte Corporation		X
Fujitsu Computer Products		X
G2 Networks		X
Gadzoox Networks		XA
GENROCO,Inc.		XA
Gordon Consultants	X	
Hewlett-Packard		X
Hitachi Computer Products (America) Inc.	X	
Hitachi Data Systems	X	X
IBM Corporation	X	X
ICP-vortex Computersysteme GmbH	X	
Infortrend Corporation	X	
Integrix,Inc.	X	
Intel Corporation		X

Intelliguard Software		X
Jems Data Unlimited, Inc.	X	
JMR Electronics, Inc.	X	
Legato Systems		X
LSI Logic Corp.	X	X
McData Corporation		X
Mercury Computer Systems		X
Meridian Data,Inc.		X
MTI Technology Corporation	X	X
Mylex Corporation	X	X
NEC Corporation	X	
Network Appliance		XA
Novell, Inc.		X
nStor Corporation	X	
ORCA Technologies	X	
Overland Data Inc.		X
Pathlight Technology,Inc.		X
Programmed Logic		X
Proware Technology Corp.	X	
QLogic		XA
Quantum		X
RAID,Inc.	X	
Seagate Technology		X
Seek Systems,Inc.	X	
Silicon Graphics		XA
SMS Data Products Group		XA
SOFTWORKS		XA
Sony		XA
Stac,Inc		XA
Sterling Software		XA
Storage Area Networks Ltd.		X
Storage Computer Corporation		X
StorageTek, Inc.	X	X
Sun Microsystems		X
Syred Data Systems	X	
Technology Forums	X	
TeraStor Corporation		X
Tricord Systems,Inc.	X	X
Trimm Technologies		XA
Unisys Corporation	X	XA
United Digital Limited	X	
Veritas Software Corporation	X	X
Vinca		XA
Vixel Corporation		X

Note: For SNIA, XA denotes associate members.

Table 3. Firms Entering the Disk Array Market by Entry Mode and Origin Industry

Total De Novo:		**45**
Total De Alio:		**212**
Computers		**45**
Integrators		**35**
Systems int.	14	
Storage int.	19	
Server int.	1	
Other int.	1	
Storage Subsytems		**27**
Subsystems	22	
Subs. Disk	2	
Subs. Optical	2	
Subs. Tape	1	
Storage (Diversified)		**19**
Value Added Resellers		**12**
Controllers		**10**
Software		**11**
Software	7	
Soft.(Storage)	4	
Diversified Electronics		**6**
Controller Cards		**5**
Distribution		**5**
Memory		**4**
Peripherals		**4**
Computers and Periph.		**4**
Adapters		**3**
Computer Add-Ons		**3**
Hard Disk Drives		**3**
Computer Upgrade		**2**
Controller and Memory		**1**
Computer Server		**1**
Computer & Storage Subs.		**1**
Computers/Disk Drive		**1**
Peripherals & Storage Sub.		**1**
Boards, Chips and Soft.		**1**
Motherboards		**1**
Servers (Video)		**1**
Storage (Optical)		**1**
Storage (RAM)		**1**
Storage Cabinets and Enc.		**1**
Ethernet Trans. and Hubs		**1**
Network Products		**1**
Diversified Other		**1**

Figure 1. Historical Summary of Developments in the Disk Drive Array Market

1956	First disk drive.
1965	First controller handling multiple drives (IBM 2314 DASF).
1966-70	Disk drives now provided storage for on-line processing, which became the dominant mode in most systems. Data reliability achieved through multiple copies on disk packs.
1971	The IBM 3330 Facility improved data integrity by extensive error detection and correction capabilities.
1976	First Tandem fault-tolerant computer shipped.
1978	IBM receives U.S. patent for "System for Recovering Data Stored in Failed Memory Unit."
Mid-1980s	First disk subsystem (multiple small form factor disk drives for higher performance and capacity storage) used for the PC environment.
1987	U.C. Berkeley "RAID" technical paper presented at conference.
1988	First shipment of a disk drive array or "cluster" using small form factor disk drives (5.25-inch or 3.5-inch).
1990	EMC offers the first disk array for mainframe storage that incorporates 5.25-inch disks. Its Symmetrix 4200 Integrated Cached Disk Array (ICDA) is a 24-gigabyte RAID system that replaces traditional 14" DASD disks.
1992	RAID Advisory Board established.
1997	Storage Networking Industry Association established.

Figure 2. Density of Producers in Disk Array Market by Year

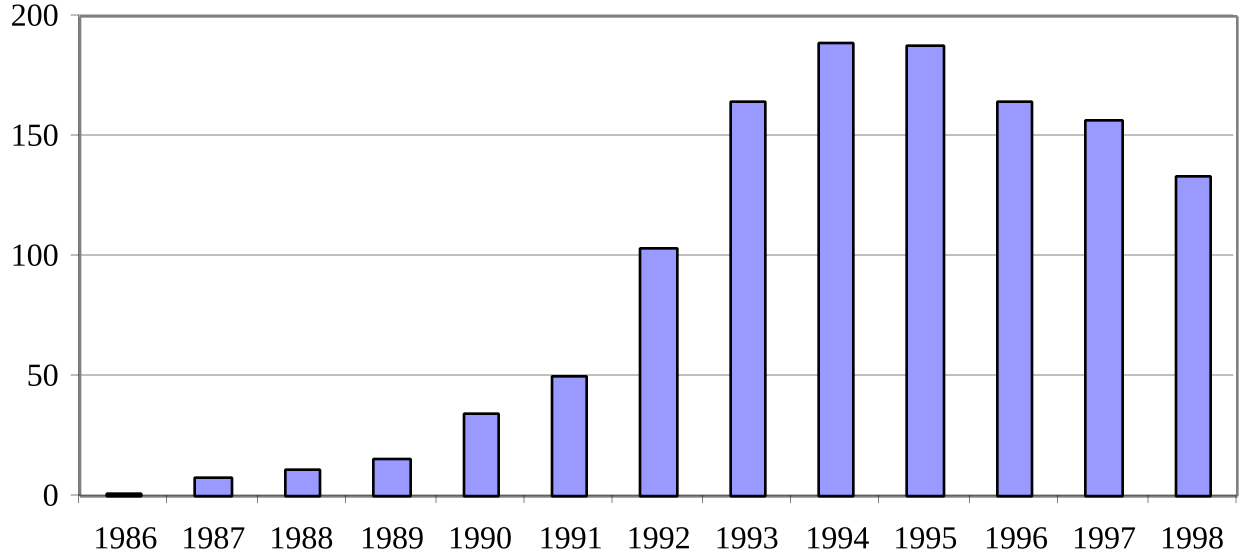


Figure 3. Entries of Producers in Disk Array Market by Year

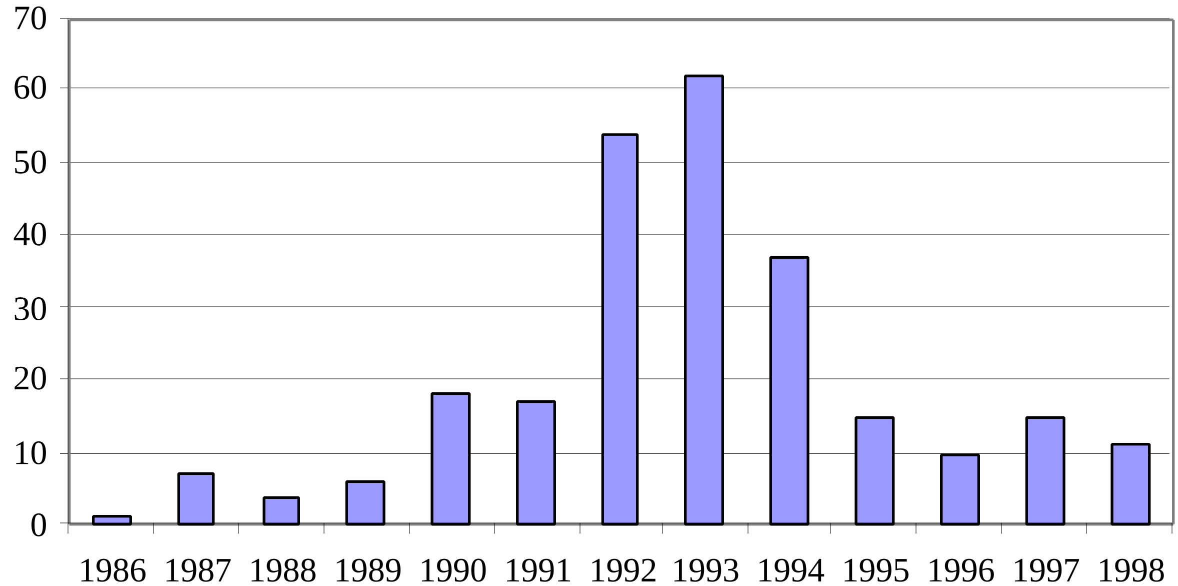


Figure 4. Exits of Producers in Disk Array Market by Year

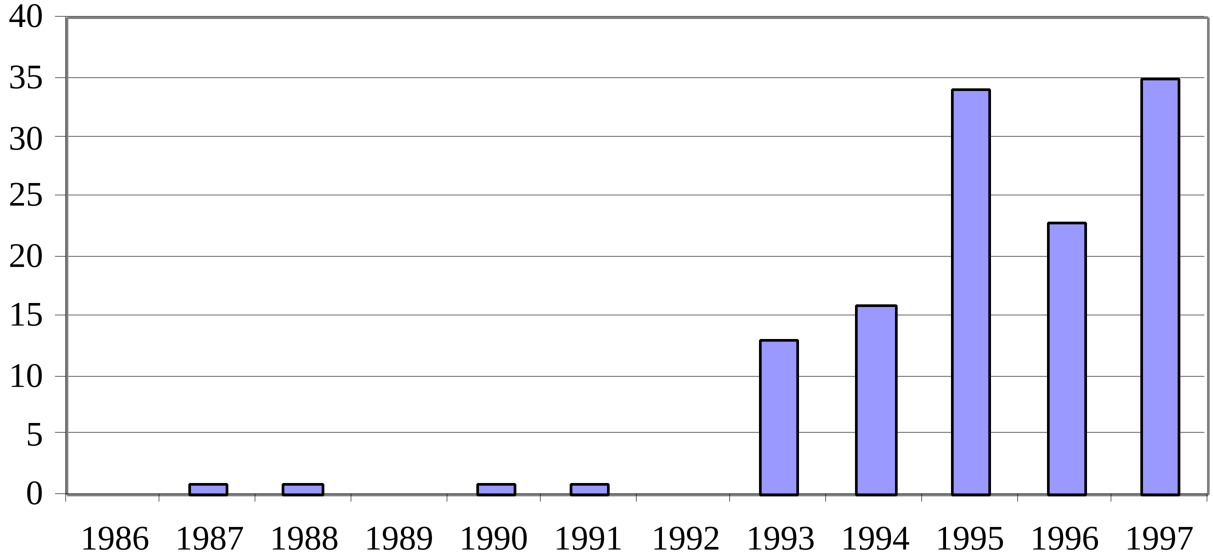


Figure 5. Density of De Novo Firms

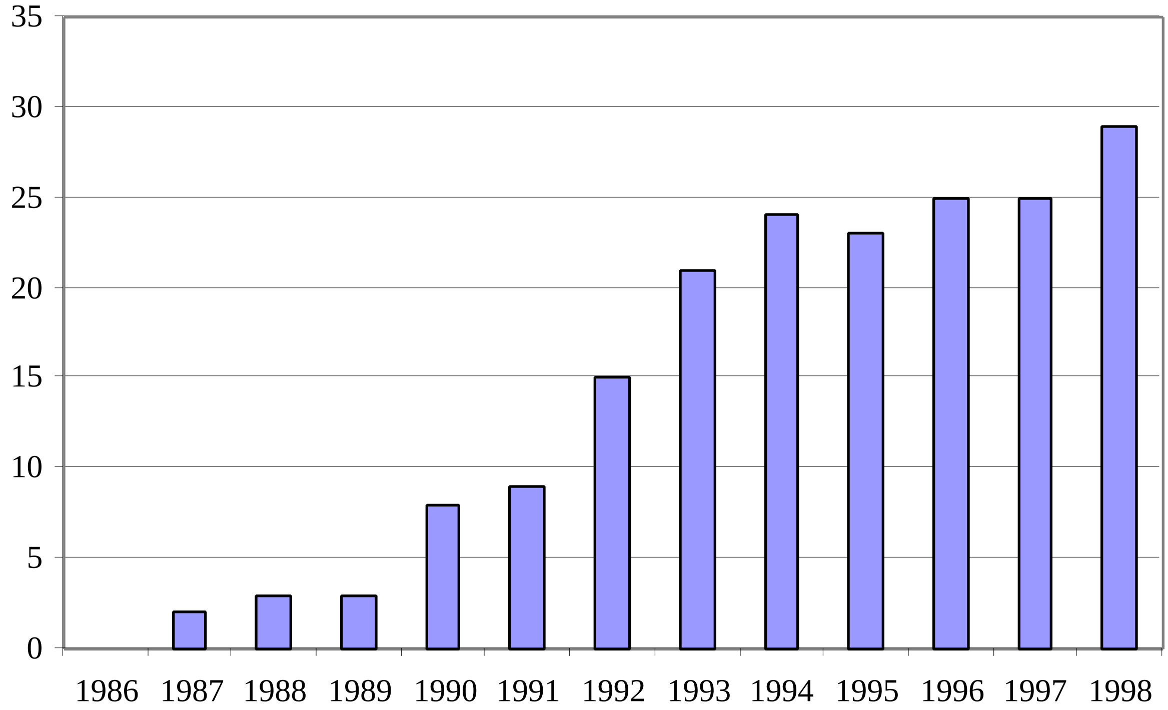
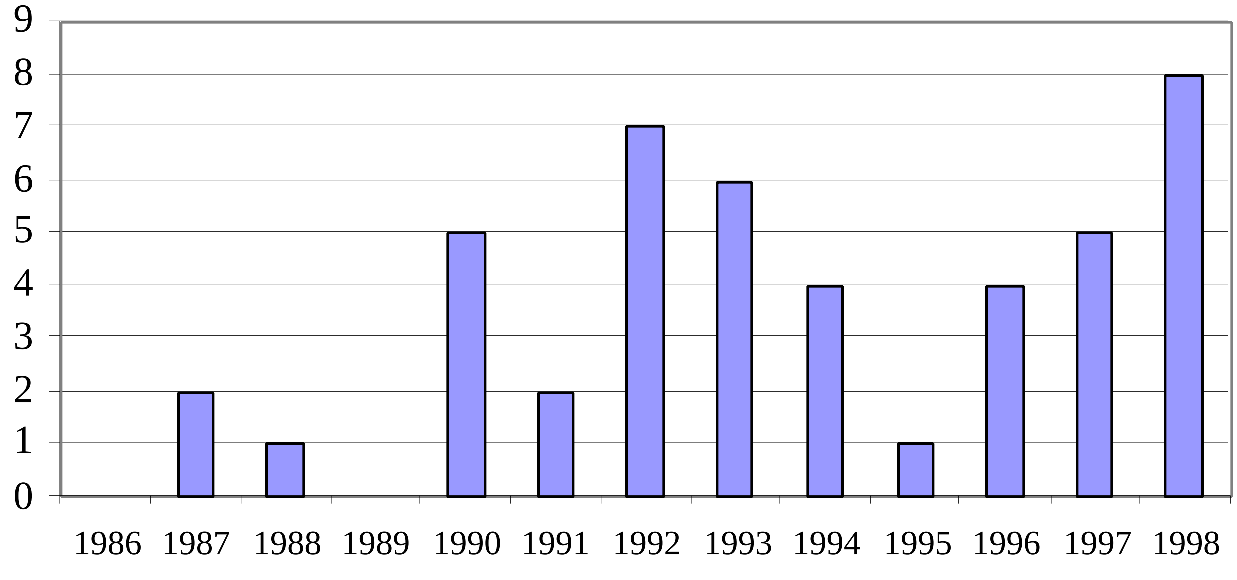


Figure 6. De Novo Entries



NOTES

[1] The theory developed by Polos et al. (1999) has the advantage of combining feature-based, network-based and boundary-based approaches to forms.

[2] Valuation is used here with respect to expectations about the form, not its social or economic value per se. To distinguish the ideas, Polos et al. (1999) refer to their concept as κ-valuation.

[3] Some very recent efforts avoid this problem by shifting the level of analysis upward, to the community level (Ruef 1999) or to environment and social movements within it (Swaminathan and Wade 1999). We applaud these efforts and think they hold great promise. However, we also believe that they will not necessarily yield the same insights as our approach, a position which we attempt to demonstrate below when we consider theory in light of the "facts" of our case. In general, our view is that it is potentially insightful to address important and difficult problems such as form emergence in a variety of ways, especially if they show an ability to rise above deficiencies of previous efforts.

[4] Thinking Machines offered a storage system called Data Vault for use with its CM-2 system, which it claimed was the world’s fastest computer (Computerworld, 1987; The Christian Science Monitor, 1990). Its array employed large form factor disk drives.

[5] RAID is actually an acronym standing for Redundant Array of Independent Disks. At some point, RAID 0, which the Berkeley team did not mention, emerged as a sixth level. But it is really a misnomer. It is striping *without* redundancy -- a pure play to maximize data transfer rates.

[6] The RAID Advisory Board was not a trade association in the legal sense. It was created and chaired by a consultant to the disk array industry, and membership fees were used to compensate the consultant as well as implement RAB’s mission, which was functionally equivalent to that of a formal association: to clarify RAID terminology, educate the market (particularly the engineering community), and promote common interests among operating system developers and standards bodies for the fledgling industry.

[7] Twincom would continue to operate through the end of the observation period (end of 1998).

[8] The example is counterfactual in more ways than one. By existing laws, the major brewers would have been legally prohibited from operating brewpubs, which are small-scale, vertically integrated operations.